Likelihood Regret: An Out-of-Distribution Detection Score For Variational Auto-encoder

Abstract

Deep probabilistic generative models enable modeling the likelihoods of very high dimensional data. An important application of generative modeling should be the ability to detect out-of-distribution (OOD) samples by setting a threshold on the likelihood. However, a recent study shows that probabilistic generative models can, in some cases, assign higher likelihoods on certain types of OOD samples, making the OOD detection rules based on likelihood threshold problematic. To address this issue, several OOD detection methods have been proposed for deep generative models. In this paper, we make the observation that some of these methods fail when applied to generative models based on Variational Auto-encoders (VAE). As an alternative, we propose Likelihood Regret, an efficient OOD score for VAEs. We benchmark our proposed method over existing approaches, and empirical results suggest that our method obtains the best overall OOD detection performances compared with other OOD method applied on VAE.